

CLAIMS

What is claimed is:

1. A speech processing method comprising:
receiving speech signals;
processing the received speech signals to generate a plurality of phoneme clusters;
grouping the phoneme clusters into a first cluster node and a second cluster node; and
determining automatically if a phoneme cluster in the first cluster node is to be moved into the second cluster node based on a likelihood increase of the phone cluster of the first cluster node from being in the first cluster node to being in the second cluster node.
2. The speech processing method of claim 1, further comprising:
moving the phoneme cluster in the first cluster node into the second cluster node if the first cluster node is determined to be moved into the second cluster node.
3. The speech processing method of claim 2, wherein moving the phoneme cluster in the first cluster node into the second cluster node includes:
moving the first cluster node into the second cluster node if the most likelihood increase is more than a threshold value.
4. The speech processing method of claim 1, wherein the phoneme clusters are triphone clusters based on a hidden markov model (HMM).
5. The method of claim 1, wherein the grouping of the phoneme clusters includes:
grouping the triphone clusters according to answers to best phonetic context based questions related to the triphone clusters.

6. A speech processing system comprising:
 - an input to receive speech signals;
 - a processing unit to process received speech signals, to generate a plurality of phoneme clusters from the processed received speech signals, to group the phoneme clusters into a first cluster node and a second cluster node, and to determine automatically if a phoneme cluster in the first cluster node is to be moved into the second cluster node based on a likelihood increase of the phone cluster of the first cluster node from being in the first cluster node to being in the second cluster node.
7. The speech processing system of claim 6, wherein the processing unit is to move the phoneme cluster in the first cluster node into the second cluster node if the first cluster node is determined to be moved into the second cluster node.
8. The speech processing system of claim 7, wherein the processing unit is to move the first cluster node into the second cluster node if the most likelihood increase is more than a threshold value.
9. The speech processing system of claim 6, wherein the phoneme clusters are triphone clusters based on a hidden markov model (HMM).
10. The speech processing system of claim 9, wherein the processing unit is to group the triphone clusters according to answers to best phonetic context based questions related to the triphone clusters.
11. A machine-readable medium that provides instructions, which if executed by a processor, cause the processor to perform the operations comprising:
 - receiving speech signals;
 - processing the received speech signals to generate a plurality of phoneme clusters;
 - grouping the phoneme clusters into a first cluster node and a second cluster node; and

determining automatically if a phoneme cluster in the first cluster node is to be moved into the second cluster node based on a likelihood increase of the phone cluster of the first cluster node from being in the first cluster node to being in the second cluster node.

12. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations comprising:

moving the phoneme cluster in the first cluster node into the second cluster node if the first cluster node is determined to be moved into the second cluster node.

13. The machine-readable medium of claim 12, further providing instructions, which if executed by a processor, cause the processor to perform the operations comprising:

moving the first cluster node into the second cluster node if the most likelihood increase is more than a threshold value.

14. The machine-readable medium of claim 11, further providing instructions, which if executed by a processor, cause the processor to perform the operations comprising:

processing the received speech signals to generate a plurality of phoneme clusters that are triphone clusters based on a hidden markov model (HMM).

15. The machine-readable medium of claim 14, further providing instructions, which if executed by a processor, cause the processor to perform the operations comprising:

grouping the triphone clusters according to answers to best phonetic context based questions related to the triphone clusters.